AFRL-SA-WP-SR-2015-0025

A Preliminary Investigation into Cognitive Aptitudes Predictive of Overall MQ-1 Predator Pilot Qualification Training Performance

Wayne Chappelle, Emily Skinner, Krista Traut
U.S. Air Force School of Aerospace Medicine, Wright-Patterson AFB, OH

Tanya Goodman, Julie Swearingen, and William Thompson
NeuroStat Analytical Solutions, LLC, San Antonio, TX

November 2015

Air Force Research Laboratory
711th Human Performance Wing
U.S. Air Force School of Aerospace Medicine
Aeromedical Research Department
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

NOTICE AND SIGNATURE PAGE

AFRL-SA-WP-SR-2015-0025 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNATURE//	//SIGNATURE//
MS. JULIA PARAKKAT Chief, Operational Health & Perform Res Div	DR. RICHARD A. HERSACK Chair, Aeromedical Research Department

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
6 Nov 2015	Special Report	September 2014 – September 2015

4. TITLE AND SUBTITLE

A Preliminary Investigation into Cognitive Aptitudes Predictive of Overall MQ-1 Predator Pilot Qualification Training Performance

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Wayne Chappelle, Tanya Goodman, Julie Swearingen, and William Thompson

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

USAF School of Aerospace Medicine
Aeromedical Research Department/FHO
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

8. PERFORMING ORGANIZATION REPORT NUMBER

AFRL-SA-WP-SR-2015-0025

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release.

13. SUPPLEMENTARY NOTES

Cleared, 88PA, Case # 2015-1652.

14. ABSTRACT

United States Air Force MQ-1 Predator remotely piloted aircraft pilots are critical assets in a challenging, high-risk, and rapidly evolving aviation career field. However, the Air Force has had difficulties over the past decade recruiting candidates with the "right stuff" and filling Predator pilot vacancies. The purpose of this study was to evaluate computer-based intelligence and neuropsychological testing on training candidates (*n* = 235) obtained during medical flight screening and prior to training that distinguishes elite vs. adequate MQ-1 qualification training performance. Findings reveal a combination of cognitive aptitudes that distinguish elite levels of training performance among those selected for this career field. Results may be utilized to improve personnel selection and aeromedical screening procedures for identifying suitable candidates for this high-risk, high-demand occupation.

15. SUBJECT TERMS

Remotely piloted aircraft, RPA, neuropsychological screening, intelligence testing, computer-based testing, Predator, MQ-1

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Wayne Chappelle
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	SAR	20	

This page intentionally left blank.

TABLE OF CONTENTS

Section **Page**

LIST OF TABLES

Page

1.0 SUMMARY

United States Air Force MQ-1 Predator remotely piloted aircraft pilots are critical assets in a challenging, high-risk, and rapidly evolving career field. However, the Air Force has had difficulties over the past decade recruiting candidates with the "right stuff" and filling Predator pilot vacancies. The purpose of this study was to evaluate computer-based intelligence and neuropsychological testing on training candidates (n = 235) obtained during medical flight screening that distinguishes elite vs. adequate MQ-1 qualification training performance. Findings reveal a combination of cognitive aptitudes that distinguish elite levels of training performance among those selected for this career field. Results may be utilized to improve personnel selection and aeromedical screening procedures for identifying suitable candidates for this high-risk, high-demand occupation.

2.0 INTRODUCTION

United States Air Force (USAF) MQ-1 Predator remotely piloted aircraft (RPA) pilots are in a high-risk profession engaged in a wide range of reconnaissance and weapons-strike operations where mistakes in the performance of their flying duties can lead to significant costs in terms of human life, national security, and foreign relations. Each year there are several aviation-related mishaps in training and operational missions where human error is considered a causal factor. As a result of the high-risk nature of flying an RPA and the evidence of a pilot's psychological disposition as crucial for safe and effective flying, the intellectual and neuropsychological assessment of USAF RPA pilots is a critical part of selection and aeromedical screening procedures.

USAF Predator RPA pilots have served a variety of roles in providing real-time information to commanders for tracking enemy movements and assets, safeguarding convoys, directing and protecting ground forces, locating and destroying weapons caches, identifying fixed and moving targets, catching insurgents planting roadside bombs, tracking and/or eliminating enemy combatants, augmenting manned-strike missions, surveying post-strike battle damage, as well as providing close air support and precision weapons strikes [1]. As a result, the global demand for Predator operations has increased significantly and military agencies are having difficulty recruiting and filling the increasing number of RPA pilot vacancies [2,3].

Although aspects of RPA operations are highly automated, such pilots perform a wide range of manual and computer-based tasks to actively and/or passively control, maneuver, and fly the aircraft during planned, unplanned, and critical in-flight situations. These tasks require pilots to engage in complex decision making, to include actively analyzing and responding to large amounts of simultaneous, real-time auditory and visual information and executing specific and complex tasks within a limited and often immediate time frame. As can be surmised, flying the MQ-1 Predator requires a high level of skill and ability as a pilot.[a]

[a] Nagy JE, Kalita SW, Eaton G. U.S. Air Force unmanned aircraft systems performance analyses, Predator pilot front end analysis (FEA) report. Brooks City-Base (TX): 311th Human Systems Wing, Performance Enhancement Directorate; 2006. Technical Report SURVIAC-TR-06-203. Available through the Defense Technical Information Center to U.S. Government agencies and their contractors only.

Comprehensive qualitative reviews of the knowledge, skills, and abilities of MQ-1 Predator pilots have identified several cognitive aptitudes as key to performance, including sustained attention and concentration, perceptual reasoning and processing, visual memory and learning, spatial and symbolic reasoning, general speed and accuracy of information processing, psychomotor dexterity, reaction time, as well as visual tracking, searching, and scanning [4].[b] Bailey speculated that cognitive aptitudes (e.g., visual perceptual reasoning) contributed to about two-thirds of the factors associated with MQ-1 Predator pilot performance.[b] The results are similar to other qualitative studies postulating on the cognitive aptitudes critical to the performance of military RPA pilots flying large-sized aircraft at medium to high altitudes across U.S. and foreign airspace [5].[a,c] However, the research studies cited above regarding cognitive predictors of performance are based on theoretical speculation and qualitative reviews.

Despite the limited number of studies with objective testing, research has found those who enter USAF Predator pilot training have high levels of pre-training cognitive abilities (intellectual and neuropsychological). In particular, two studies revealed MQ-1 Predator RPA pilot training candidates to have high average to superior levels of general intellectual functioning, speed and accuracy of information processing, visual reasoning and memory, psychomotor reaction time, as well as attention and concentration when compared with peers in the general normative population [4,6]. These findings are similar to the cognitive abilities of USAF pilots who fly manned airframes [7-9].

However, existing literature on the general intelligence and cognitive aptitudes specific to (and predictive of) performance outcomes of USAF Predator pilots remains limited and controversial. Anecdotal discussions with RPA Predator line and training leadership reveal a wide range of divergent opinions regarding cognitive aptitudes that distinguish those with the highest levels of overall training performance. For instance, some line commanders have postulated that faster psychomotor reaction time and information processing speed, as well as high levels of memory, distinguish such performers, whereas others have postulated higher levels of attention, spatial analytical abilities, and general cognitive functioning are the key aptitudes that delineate those with the highest overall performance.

However, identifying the cognitive aptitudes that delineate (and are predictive of) those with overall elite training performance may be particularly challenging given the entire pool of training candidates has high average to superiors levels of performance across a range of cognitive aptitudes measured via tests of intelligence and neuropsychological functioning [4,6]. Furthermore, RPA candidates are selected based on their performance on measures of general mental ability (i.e., Armed Forces Officer Qualifying Test, Test of Basic Attributes) and academic achievement. As a result, it is likely that any study attempting to identify pre-training cognitive aptitudes that delineate those with the highest levels of overall training performance will be subjected to range restriction.

---

[b] Bailey M. Predator pilot and sensor operator selection test batteries. United Kingdom: Cranwell Royal Air Force Base; 2009. Royal Air Force Technical Report. Available by request only.

[c] Nagy JE, Muse K, Eaton G, Phillips A. U.S. Air Force unmanned aircraft systems performance analyses: Global Hawk pilot and sensor operator front end analysis (FEA) report. Brooks City-Base (TX): 311th Human Systems Wing, Performance Enhancement Directorate; 2007. Technical Report SURVIAC-TR-10-041. Available through the Defense Technical Information Center to U.S. Government agencies and their contractors only.

Nonetheless, identifying the pre-training aptitudes that distinguish and predict those with elite levels of overall training performance may provide new insights for improving selection and aeromedical screening processes, as well as training outcomes. This is especially important given the rising dependency upon RPA weapons bearing aircraft for supporting multiple and continuous surveillance, reconnaissance, and combat weapons-strike operations throughout the globe and the importance of recruiting and filling the increasing number of MQ-1 Predator RPA pilot vacancies affecting readiness levels of units across the USAF [2,3].

The purpose of this exploratory study is to (a) assess differences among candidates grouped according to elite versus adequate training performance on measures of intelligence and neuropsychological testing and (b) identify cognitive aptitudes predictive of performance outcomes. It is hypothesized those with elite performance will have higher levels of cognitive functioning in the areas of visual-perceptual functioning (i.e., visual reasoning, visual memory, and learning) and speed of information processing due to the visually complex and demanding nature of flying the MQ-1 Predator and the results of comprehensive qualitative reviews [4,5][b] emphasizing the importance of such abilities.

3.0 METHODS

3.1 Participants

In total, 89 participants were selected from a pool of 235 USAF MQ-1 Predator pilot training candidates who completed Initial Qualification Training (IQT). Participants were selected based upon their average daily scores throughout training. High performers were identified as participants with scores in the upper quintile (i.e., top 20%) of those who completed training. Adequate performers were identified as participants with scores in the lowest quintile (bottom 20%) of those who completed training.

In total, 46 out of 235 (19.57%) pilots who completed IQT were identified as the highest (upper quintile) in overall training performance. This group consisted of 44 males (95.65%) and 2 females (4.35%). A total of 42 high performers (91.30%) reported Caucasian as their ethnicity/race, and 1 high performer (2.17%) reported Asian as his ethnicity/race. The average age for this group was 21.98 years (standard deviation (*SD*) = 2.62), with a range of 19 – 32 years. Data were not available for ethnicity/race on three elite performers (6.52%) and for age on two high performers (4.35%).

Overall, 43 out of 235 (18.30%) pilots who completed IQT were identified with adequate performance (lower quintile). This group consisted of 40 males (93.02%) and 3 females (6.98%). A total of 34 adequate performers (79.07%) reported Caucasian as their ethnicity/race, 2 (4.65%) reported Asian, 2 (4.65%) reported African, and 1 (2.32%) reported Hispanic. The average age for this group was 23.86 years (*SD* = 2.98), with a range of 19 – 32 years. Data were not available for ethnicity/race on four adequate performers (9.30%) and for age on seven adequate performers (16.28%).

3.2 Measures

3.2.1 Intelligence Testing. The Multidimensional Aptitude Battery II (MAB-II) assesses general intelligence with five distinct verbal intelligence subscales, five distinct performance

intelligence subscales, a verbal intelligence quotient, a performance intelligence quotient, and an overall general intelligence quotient [10].

The MAB-II is a computer-based test for administration and scoring and is a part of baseline testing during medical flight screening for all USAF RPA pilots [11]. The measure is based upon the structure and content of the Wechsler Adult Intelligence Scale and Wechsler's theory of intelligence [12]. The MAB-II manual has well-documented internal consistency reliability of 0.94 – 0.97 on the subscales and 0.92 – 0.95 on the intelligence quotients and test-retest reliability ranging from 0.83 – 0.97 on the subscales and 0.94 – 0.97 on the intelligence quotients [10]. For the general population normative scores, the standardized scores for the MAB-II subscales have a mean of 50 (SD = 10), the intelligence quotients have a mean of 100 (SD = 15), and they have been statistically corrected for age [10].

3.2.2 Neuropsychological Screening. The MicroCog was used to assess neuropsychological functioning [13]. Eighteen subtests are combined to yield five first-level indices representative of five neuropsychological domains: Attention/Mental Control, Memory, Reasoning/Calculation, Spatial Processing, and Reaction Time. Two second-level indices of Information Processing Speed and Information Processing Accuracy are also derived from the subtests but take into account overall speed and accuracy rather than specific domains. Two third-level indices of General Cognitive Functioning and General Cognitive Proficiency are also assessed. General Cognitive Functioning is derived from the two second-level indices and General Cognitive Proficiency is derived from the individual subtest proficiency scores, including both information processing speed and accuracy with weighting on accuracy of processing [13].

The MicroCog is a computer-based test for administration and scoring and is also a part of medical baseline testing during medical flight screening for all USAF RPA pilots [6,11]. Multiple-choice items with varying numbers of response options and free-response items are presented. MicroCog normative scores for the five first-level indices, two second-level indices, and two third-level indices each have a mean of 100 (SD = 15) and have been statistically adjusted for age and level of education (bachelor's degree in the current study). The manual for the MicroCog documents internal consistency reliability ranging from 0.83 – 0.95 and test-retest reliability ranging from 0.73 – 0.99 on the indices for the standard form [13,14].

3.2.3 Training Outcomes. Training outcomes from the IQT course for the MQ-1 Predator aircraft were used in this study. At the time of this study, the IQT course consisted of 45 training days, with 21 ground training days and 24 flying training days, for a total of 174.5 training hours. RPA pilot trainees were rated throughout each simulator and live flight sortie (20 sorties total) using individual mission grade sheets that assess several areas of functioning: mission planning/preparation, administrative checklist procedures, display interpretation, display/data manipulation, pedestal functions, system operations, aircraft control, navigation procedures, communication and coordination procedures, general knowledge, safety, discipline, emergency mission procedures and knowledge, etc.[d] The RPA pilot trainee's performance was evaluated in a checklist format on the mission grade sheet across multiple areas of functioning and an overall performance score for that sortie was rated on a scale of 0 to 4:

[d] Air Combat Command. MQ-1B qualification training [Unpublished technical report]. Langley AFB (VA): Department of the Air Force, Headquarters Air Combat Command; 2008.

- 0 (lack of ability or knowledge; dangerous/unsafe performance)
- 1 (performance is safe, but proficiency is limited; makes errors of omission or commission)
- 2 (performance is essentially correct; recognizes and corrects errors)
- 3 (performance is correct, efficient, skillful, and without hesitation)
- 4 (performance reflects an unusually high degree of ability)

The training candidates must achieve a proficiency score of 2 or greater on each of the elements to successfully complete the course. Training outcomes for the MQ-1 IQT course consisted of averaged scores from individual mission grade sheets from 20 simulator and live flight sorties. Each sortie was graded by one instructor and each candidate was graded by multiple instructors throughout his or her training.

3.3 Procedure

The MQ-1 Predator pilots in this study were administered the MAB-II and MicroCog as a routine part of medical flight screening prior to attending training. Test administration occurred at two locations (USAF School of Aerospace Medicine and the USAF Academy). All candidates were tested under a standardized set of instructions and conditions. Training candidates were informed of the potential uses of their baseline testing for medical and research purposes. Each pilot's test scores were archived and combined later with training numerical rating outcome grades and a "rack and stack" list assessing how each pilot performed relative to all other pilots after completion of IQT.

3.4 Data Analysis

The analyses in this study were generated using SAS® software, version 9.3 (SAS Institute, Cary, NC). The dataset consisted of 235 MQ-1 Predator pilot training candidates who passed IQT training. Participants were sorted in descending order on their average MQ-1 IQT daily score and dichotomized into two groups based on a cut-off representing a change in score closest to the lowest and highest quintiles.

Descriptive statistics were computed for the high and adequate performance groups. Correlations were run on the MAB-II and MicroCog variables with the outcome variable (high vs. adequate performance).

To determine which aptitudes were predictive of high performers, stepwise binary logistic regressions were used to retain predictor effects titrating from $p < .05$ to $p < .30$, by incremental units of .05 [15]. While hierarchical logistic regression utilizing all predictors is preferred in the literature, the current study was limited by the small sample sizes in each group. While the standard requirement is 10 records per outcome variable for logistic regressions [15], research has shown that less records per outcome are possible if the outcome groups are equal [16]. It has been recommended that as low as five records per outcome variable are adequate when interpreted with caution [17]. With the smallest group ($n = 43$), five records per predictor allowed for 8.6 predictors. Parameters to $p < .50$ have been found in the literature in small sample sizes [15], but the model with $p < .30$ was used to account for the maximum number of predictors without over fitting the model.

Logistic regressions were run individually using the 10 MAB-II subscales, then for the 5 MicroCog first-level and 2 second-level indices, and finally for the MAB-II and MicroCog combined. The higher level composite scores were not used in the predictive models because they are formulated from the lower level measures (i.e., MAB-II intelligence quotients and MicroCog third-level indices). Selection of variables to remain in the model was based upon the relative impact of each variable to the increased incremental validity of the model. This was assessed by reviewing the R^2 at each step of the model, sorting the absolute value of the standardized beta estimates and the results of the receiver operating curve estimates. Goodness-of-fit statistics were performed for each model. Negative and positive predictive value rates were calculated from the logistic regression leave-one-out cross-validation classification tables.

Multivariate range restriction was also considered. With a sample consisting of only pilot training candidates who passed training (and not including those who failed training), indirect range restriction regarding performance on measures of cognitive aptitude is assumed to exist due to selection requirements [18]. The effects of direct range restriction were also noted, because this study truncated the dataset from the original (n = 235) to only the top and bottom quintiles. A comparison of the predictor intercorrelations and predictor correlations with the outcome variable for the original dataset (splitting the dataset at the 50th percentile) and the truncated dataset (top and bottom quintiles) was performed. When ranking the correlations of the predictors to the outcome variable in descending order, the order of predictors did not change when comparing the results of the original dataset and the truncated dataset. This indicates that while corrections for direct range restriction may increase the correlations with the outcome variable, the same predictors would have emerged from a corrected model because the predictors with the top relative contribution to the model would have been the same. Additionally, effects of range enhancement, or choosing only groups from the top and bottom quintiles, may have resulted in an increase in correlations with the outcome variable when compared to correlations in the normative population. Guidelines for the minimum n required for indirect range restriction were used, based on required power, restricted predictor validity, predictor intercorrelations, unrestricted predictor validity, selection ratio, and criterion reliability [19]. Accounting for the selection ratio in the truncated dataset (selection ratio of .40 to account for the two sets of quintiles) and the criterion reliability of 0.63, the minimum required sample size was 129. Additionally, this required sample size may still be underestimated, based on corrections to the formula by Raju, Edwards, and LoVerde [20]. With the final sample size in the current study of 89 not meeting the minimum required sample size, logistic regressions utilizing multivariate range restriction corrections are not reported. Multivariate range restriction correlation adjustments were calculated [21], and subsequent logistic regression analyses were conducted on uncorrected correlations. Corrections for dichotomizing the criteria were also considered, but are not reported because of the small sample size [22].

4.0 RESULTS

It is important to note all participants in this study completed MQ-1 Predator pilot training before being assigned to their operational unit for combat mission readiness training. All pilots were deemed proficient and met qualifications to graduate training. The group average training score for the elite performers, taken as an average of daily individual grade sheets, was 2.19 (SD = 0.10), with a range of 2.07 – 2.53. The group average training score for the adequate performers was 1.76 (SD = 0.11), with a range of 1.50 – 1.86. While adequate performance

represented safe, effective flying with limited proficiency, elite performance represented safe, effective flying along with proficient corrections of errors and high levels of skill.

Tables 1 and 2 provide descriptive statistics on the MAB-II and MicroCog and correlations with the outcome variable (high vs. adequate performance). Both adequate and elite performers scored at least one standard deviation higher than the age-corrected normative population (mean (*M*) = 100, *SD* = 15) on the three intelligence quotients measured by the MAB-II. In addition, elite performers scored at least one standard deviation higher than the normative population (*M* = 50, *SD* = 10) on all five verbal subscales and on four performance subscales. Adequate performers also scored at least one standard deviation higher than the normative population on one verbal subscale and three performance subscales. Standard deviations for both groups on the intelligence quotients and subscales were smaller than the normative population, showing less variance in scores for each group. Correlations with elite versus adequate performance were significant at $p < .01$ for Verbal IQ; the Information, Similarities, and Vocabulary verbal subscales; and the Digit Symbol performance subscale. A correlation was significant at $p < .05$ for the Arithmetic verbal subscale.

Table 1. MAB-II Descriptive Statistics and Correlations with Outcome (High vs. Adequate Performance)

Quotients and Subscales	High Performers *M* (*SD*) (n = 46)	Adequate Performers *M* (*SD*) (n = 43)	r	r_c
Intelligence Quotients				
Full Scale	120.26 (16.83)	118.30 (6.60)	.08	
Verbal	121.65 (5.34)	116.33 (5.55)	.44[a]	
Performance	120.17 (7.39)	117.88 (9.47)	.14	
Verbal Subscales				
Information	69.13 (4.36)	64.81 (5.70)	.40[a]	.66
Comprehension	60.33 (3.88)	58.77 (4.85)	.18	.60
Arithmetic	61.07 (7.26)	57.95 (7.18)	.21[b]	.48
Similarities	61.48 (3.80)	58.81 (4.51)	.31[a]	.57
Vocabulary	62.17 (6.49)	57.30 (6.34)	.36[a]	.60
Performance Subscales				
Digit Symbol	68.04 (6.07)	63.21 (9.12)	.30[a]	.54
Picture Completion	60.91 (6.32)	60.44 (6.46)	.04	.19
Spatial Analyses	60.72 (6.96)	59.12 (7.03)	.12	.36
Picture Arrangement	50.72 (7.35)	50.53 (8.04)	.01	.21
Object Assembly	62.24 (5.13)	61.65 (5.40)	.06	.32

Note: Column headings indicate observed correlations (r) and correlations corrected for range restriction (r_c).
[a]$p < .01$.
[b]$p < .05$. 95% CI = 95% confidence interval. Scores are corrected for age.

Both adequate and elite performers had mean group scores within one standard deviation of the age- and education-corrected normative population (M = 100, SD = 15) [13] on each of the neuropsychological indices measured by the MicroCog. Correlations with high vs. adequate performance were significant at $p < .01$ for the Memory first-level index, Information Processing Speed second-level index, and General Cognitive Proficiency third-level index. Correlations significant at $p < .05$ were the Attention/Mental Control and Reasoning/Calculation first-level indices and the General Cognitive Functioning third-level index.

Stepwise logistic regressions were run using the 10 MAB-II subscales, the 5 MicroCog first-level and 2 second-level indices, and then the MAB-II and MicroCog measures combined. Table 3 shows the final predictors for each model, Wald coefficients, and standard estimates. The leave-one-out cross-validation classification table for the MAB-II subscales stepwise logistic regression revealed 71.74% of elite performers and 55.81% of adequate performers were correctly classified. The negative predictive value was 64.86% and the positive predictive value was 63.46%. The model had acceptable goodness-of-fit when assessing area under the curve (AUC) = 0.78. Hosmer and Lemeshow $\chi^2(8) = 4.06$, $p = .85$ demonstrated acceptable goodness-of-fit. The stepwise logistic regression for the MicroCog first- and second-level indices revealed 76.09% of elite performers and 53.49% of adequate performers were correctly classified. The negative predictive value was 67.65% and the positive predictive value was 63.64%. The model had acceptable goodness-of-fit when assessing AUC = 0.76 and Hosmer and Lemeshow $\chi^2(8) = 6.02$, $p = .65$.

Table 2. MicroCog Descriptive Statistics and Correlations with Outcome (High vs. Adequate Performance)

Indices	High Performers Mean (SD) (n = 46)	Adequate Performers Mean (SD) (n = 43)	r	r_c
First Level				
Attention/Mental Control	105.52 (11.62)	99.12 (11.44)	.27[a]	.30
Reasoning/Calculation	99.04 (11.83)	94.19 (11.03)	.21[a]	.26
Memory	113.93 (15.56)	104.51 (14.23)	.30[b]	.29
Spatial Processing	107.61 (8.53)	104.02 (12.88)	.17	.25
Reaction Time	99.00 (13.61)	99.00 (13.43)	.00	.00
Second Level (Information)				
Processing Speed	109.20 (10.26)	97.42 (14.17)	.44[b]	.44
Processing Accuracy	98.67 (14.67)	98.77 (12.34)	-.00	-.07
Third Level (General)				
Cognitive Functioning	112.09 (14.03)	105.95 (14.45)	.21[a]	
Cognitive Proficiency	106.91 (9.83)	100.44 (9.62)	.32[b]	

Note: Column headings indicate observed correlations (r) and correlations corrected for range restriction (r_c). Scores are corrected for age and education.
[a]$p < .05$.
[b]$p < .01$.

Table 3. High vs. Adequate Overall MQ-1 Predator Training Performance Logistic Regression Model

Step	R^2	Predictors	Wald	*p*	Std Estimate
1	.324	MAB-II			
		Information	13.99	.00	0.32
		Vocabulary	2.55	.11	0.32
		Digit Symbol	3.49	.06	0.30
		Arithmetic	1.87	.17	0.26
		Spatial Analyses	1.43	.23	0.19
2	.291	MicroCog			
		Information Processing Speed	16.88	.00	0.55
		Memory	2.99	.08	0.24
3	.548	MAB-II & MicroCog			
		Information Processing Speed	16.88	.00	1.07
		Vocabulary	13.64	.00	0.52
		Spatial Analyses	3.53	.06	0.40
		Arithmetic	2.48	.12	0.40
		Digit Symbol	2.37	.12	0.37
		Reasoning/Calculation	2.22	.14	0.34
		Information	2.91	.09	0.33

When the MicroCog was added to the MAB-II, an additional 25.72% of variance was explained. With a criterion cut-off score of .5, the stepwise logistic regression for the combined model revealed 71.74% of elite performers (33 out of 46) and 62.79% of adequate performers (27 out of 43) were correctly classified. Positive predictive value is an assessment of how well the model successfully predicted true positives from all those predicted to be true positives. If the 16 false positives (adequate performers who were predicted to be elite performers) are added to the 33 correctly classified elite performers, there are 49 predicted elite performers. The positive predictive value for elite performers equals 67.35%, or the number of elite performers correctly classified by the total predicted to be elite performers. The negative predictive value was 67.50%. While the negative predictive value was similar to the MicroCog only model, the positive predictive value for the combined model increased when compared to both the MAB-II only and MicroCog only models. While the Hosmer and Lemeshow goodness-of-fit statistic, $\chi^2(8) = 9.67$, $p = .29$, was low, stability of the predictive model was shown by acceptable AUC = 0.88 and a less than 10% shrinkage in overall classification ability from the initial to the cross-validated leave-one-out classification.

5.0 DISCUSSION

It is important to note all MQ-1 Predator pilots in this study completed MQ-1 Predator pilot training (which occurs before being assigned to their operational unit for combat mission readiness training,) were deemed proficient in the pilot skills, and met qualifications to graduate training. While adequate performance represented safe, effective flying with limited proficiency, elite performance represented safe, effective flying along with proficient corrections of errors and high levels of skill.

The results revealed both elite and adequate performers scored higher than the general civilian normative population on intelligence testing with less variance. This resulted in an attenuated distribution (e.g., range restriction) of intelligence scores among Predator pilots from both groups when compared with the normative population. However, the scores from both groups were similar to USAF pilots of manned airframes [7,9] and to a previous study on MQ-1 Predator RPA pilot training candidates [6]. Additionally, those with elite training performance scored higher on all five of the verbal ability subscales (e.g., Information, Comprehension, Arithmetic, Similarities, and Vocabulary) and on only one out of five performance-based subscales (i.e., Digit Symbol) when compared to those with adequate performance scores. As a group, those with elite training performance had higher levels of intelligence in the areas of general knowledge, social reasoning and comprehension, numerical processing and reasoning, general conceptual and verbal reasoning, as well as visual learning and processing speed.

In regard to neuropsychological testing, there also was less variance in the scores between both groups of pilots when compared with peers in the general civilian normative population. The scores from both groups were similar to USAF manned airframe pilots [7,9]. However, as a group, pilots with elite training performance also scored higher on neuropsychological measures of attention and concentration, reasoning, memory, spatial processing, as well as speed of information processing. These areas of neuropsychological functioning were postulated as critical to adaptation to Predator pilot job skills and demands in an earlier qualitative study involving interviews with subject matter experts [4].

5.1 Intelligence and Neuropsychological Predictors of High vs. Adequate Performance

The findings of Information Processing Speed and Vocabulary (a measure of verbal ability) as significant predictors of performance suggest those within the elite performance group had a higher capacity to effectively process and respond to multiple inputs of visual and auditory information with a more robust set of vocabulary skills to draw from for effectively articulating oneself and communicating with others in regard to information, tasks, and procedures. Such aptitudes are reasonably perceived as essential for minimizing the risk for mishaps and responding to both routine and emergent (and often unpredictable) conditions during flight.

The findings of Spatial Analyses, Arithmetic, and Digit Symbol measures as significant predictors of performance suggest those in the elite performance group had higher visual-spatial reasoning, numerical processing, and working memory aptitudes and are consistent with previous research identifying such aptitudes as an area of functioning distinguishing USAF pilot candidates [11]. Such aptitudes are reasonably perceived to influence how quickly and accurately a training candidate can calculate the position and distance of the aircraft between fixed and moving targets, as well as calculating, understanding, and remembering spatial relations in general. These aptitudes also appear particularly important for effectively responding to the visually complex and highly demanding spatial and numerical calculation tasks during flight, as well as the construction of mental representations of object configurations from images on several screens representing constantly changing geographical images and numerical data.

The finding of the neuropsychological Reasoning and Calculation index as predictive of performance suggests those with elite performance had higher capacity for concept formation, cognitive flexibility, inductive reasoning, and performing mental calculations. Such aptitudes also represent a higher capacity for critical thinking and problem solving within a condensed,

limited period of time. Processing and identifying information in a logical fashion and effectively labelling, prioritizing, and attending to important information when inundated with large amounts of visual and auditory data are key to performance. This is particularly important to analyzing various facts and points of view and quickly identifying errors or areas of concern related to specific flying tasks.

The finding of the Information subscale (general fund of knowledge for world-related facts and information) as a significant predictor of performance was unexpected given that general fund of knowledge is often influenced by educational level and all USAF pilots are college graduates. However, further analyses of this standing suggest those with elite performance have a higher capacity for storing and retrieving information in general. This may be particularly important when learning new tasks and procedures and forming an understanding of conditions as they emerge.

Although intelligence testing alone explained 29.11% of the variance in performance, the addition of neuropsychological testing increased the variance to 54.83%. Of the elite performers, the model correctly classified 71.74% as elite performers. The results suggest that a collective combination of intelligence and neuropsychological testing tapping into speed of information processing, visual-spatial analyses, reasoning, numerical processing, visual memory, and vocabulary may help to improve aeromedical screening and personnel classification processes for identifying recruits with the "right stuff" for MQ-1 Predator RPA pilot training.

5.2 Limitations

This study had limitations that warrant discussion. First, the group sample sizes for both the adequate and high performance groups were small (*less than* 50). This may have hindered the ability to perform multivariate range restriction corrections. Performing range restriction adjustments on small sample sizes may lead to results with a 50% chance or less of detecting validity, if validity exists. Second, caution should be given to generalizing the results from this study to other RPA pilot career fields within civilian and military industries. Performance and task requirements for other RPA airframes (such as low-altitude, non-weapons-bearing platforms) may differ significantly. Third, repeated studies are needed to assess for differences regarding minority group status. Studies evaluating cognitive performance should remain sensitive to differences between minority groups to ensure objective and unbiased decision making when utilizing the study results for modifying selection and aeromedical screening processes. Fourth, the inter-rater reliability between instructors on the average daily training scores was not accessed because the data in the current study included an averaged score and not daily scores or information on the grading instructor. Additional studies are needed to replicate these findings and provide more confidence to the current results.

5.3 Conclusions

MQ-1 Predator RPA pilots are required to perform complex decision making, analyze real-time auditory and visual information, and understand and perform effective task prioritization in a high-demand, 24-hour operational environment. The results of this study reveal cognitive and neuropsychological aptitudes have a key role in performance outcomes for those pursuing this unique aviation-related career field. The results suggest that including the assessment of key areas of psychological functioning identified in this study may help to improve personnel selection as well as aeromedical screening procedures.

6.0 REFERENCES

1. U.S. Air Force. The U.S. Air Force remotely piloted aircraft and unmanned aerial vehicle strategic vision. Washington (DC): Department of the Air Force; 2005. Research Paper 1. [Accessed Retrieved from 1 Dec 2014]. Available from http://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1000&context=usafresearch .
2. Hoagland BT. Manning the next unmanned Air Force: developing RPA pilots of the future. Washington (DC): Center for 21st Century Security and Intelligence, The Brookings Institute; 2013. [Accessed 1 Dec 2014]. Available from http://www.brookings.edu/research/papers/2013/08/06-air-force-drone-pilot-development-hoagland.
3. Government Accountability Office. Air Force: actions needed to strengthen management of unmanned aerial system pilots. Washington (DC): GAO; 2014. Report GAO-14-316.
4. Chappelle W, McDonald K, McMillan K. Important and critical psychological attributes of USAF MQ-1 Predator and MQ-9 Reaper pilots according to subject matter experts. Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine; 2011. Technical Report AFRL-SA-WP-TR-2011-0002.
5. Paullin C, Ingerick M, Trippe MD, Wasko L. Identifying best bet entry-level selection measures for US Air Force remotely piloted aircraft (RPA) pilot and sensor operator (SO) occupations. Randolph AFB (TX): Air Force Center Personnel Center; 2011. Technical Report AFCAPS-FR-2011-0013. [Accessed 1 Dec 2014]. Available from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA554209.
6. Chappelle W, Tran NVT, Thompson W, Goodman T, Hyde K. Intelligence and neuropsychological aptitude testing of U.S. Air Force MQ-1 Predator pilot training candidates. Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine; 2012. Technical Report AFRL-SA-WP-TR-2013-0003.
7. Chappelle W, Ree MJ, Barto EL, Teachout MS, Thompson WT. Joint use of the MAB-II and MicroCog for improvements in the clinical and neuropsychological screening and aeromedical waiver process of rated USAF pilots. Brooks City-Base (TX): U.S. Air Force School of Aerospace Medicine; 2010. Technical Report AFRL-SA-BR-TR-2010-0002.
8. King RE, Barto E, Ree MJ, Teachout MS, Retzlaff PD. Compilation of pilot cognitive ability norms. Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine; 2011. Technical Report AFRL-SA-WP-TR-2012-0001.
9. King RE, Carretta TR, Retzlaff P, Barto E, Ree MJ, Teachout MS. Standard cognitive psychological tests predict military pilot training outcomes. Aviation Psychology and Applied Human Factors. 2013; 3(1):28-38.

10. Jackson DN. Multidimensional Aptitude Battery-II: manual. Port Huron (MI): Sigma Assessment Systems, Inc.; 2003.
11. Chappelle W, Heerema BD, Thompson WT. Factor analysis of computer-based Multidimensional Aptitude Battery-Second Edition intelligence testing from rated U.S. Air Force pilots. Wright-Patterson AFB (OH): U.S. Air Force School of Aerospace Medicine; 2012. Technical Report AFRL-SA-WP-TR-2013-0005.
12. Wechsler D. Wechsler Adult Intelligence Scale-3rd Edition (WAIS-3®). San Antonio (TX): Harcourt Assessment; 1997.
13. Powell DH, Kaplan EF, Whitla D, Weintraub S, Caitlin R, Funkenstein HH. MicroCog: assessment of cognitive functioning (version 2.1) manual. San Antonio (TX): Psychological Corporation, Harcourt Brace & Co.; 1993.
14. Powell DH. Profiles in cognitive aging. Cambridge (MA): Harvard University Press; 1994.
15. Hosmer DW, Lemeshow S. Applied logistic regression, 2nd ed. New York (NY): John Wiley & Sons, Inc.; 2000.
16. Spicer J. Making sense of multivariate data analysis. Thousand Oaks (CA): Sage Publications; 2005.
17. Vittinghoff E, McCulloch CE. Relaxing the rule of ten events per variable in logistic and Cox regression. Am J Epidemiol. 2007; 165(6):710-718.
18. Carretta TR. Pilot candidate selection method: still an effective predictor of US Air Force pilot training performance. Aviation Psychology and Applied Human Factors. 2011; 1(1):3-8.
19. Sackett PR, Wade BE. On the feasibility of criterion-related validity: the effects of range restriction assumptions on needed sample size. J Appl Psychol. 1983; 68(3):374-381.
20. Raju NS, Edwards JE, LoVerde MA. Corrected formulas for computing sample sizes under indirect range restriction. J Appl Psychol. 1985; 70(3):565-566.
21. Lawley DN. IV.—A note on Karl Pearson's selection formulae. Proc R Soc Edinb A Math Phys Sci. 1944; 62(1):28-30. doi: 10.1017/S0080454100006385
22. Cohen J. The cost of dichotomization. Appl Psychol Meas. 1983; 7(3):249-253.

LIST OF ABBREVIATIONS AND ACRONYMS

AUC	area under the curve
IQT	Initial Qualification Training
M	mean
MAB-II	Multidimensional Aptitude Battery II
RPA	remotely piloted aircraft
SD	standard deviation
USAF	United States Air Force